

Common problems of mobile applications for foreign language testing

Jesús García Laborda
Dept. of Modern Philology
Universidad de Alcalá
Madrid, Spain
jesusgalaborda@wanadoo.es

Teresa Magal-Royo
DEGI
Universitat Politècnica de Valencia
Valencia, Spain
tmagal@degi.upv

José Luis Giménez López
DEGI
Universitat Politècnica de Valencia
Valencia, Spain

Abstract—As the use of mobile learning educational applications has become more common anywhere in the world, new concerns have appeared in the classroom, human interaction in software engineering and ergonomics. New tests of foreign languages for a number of purposes have become more and more common recently. However, studies interrelating language tests and ergonomics have lagged behind to the point that there is a clear lack of balance between software for m-learning and the capacities of modern equipment. This paper is based on the experience acquired through the use of mobile phones emulators for language testing. The paper does not deal with the experimental phase itself but suggests the constraints found in such experimentation from a descriptive perspective.

Keywords—*m-learning, foreign languages, testing, visual ergonomics*

I. INTRODUCTION

Many users would agree that, had it not been for the increasing number of mobile applications such as cell phones, or game platforms, their capacity of using their time constructively while commuting or in their free time might never have occurred. This is a direct result of the development of two aspects: appropriate and powerful mobile tools and also their capacity to engage the users into constructivist knowledge actions. However, most users' efforts go with endless games to fill their idle time. Only recently, intelligence development games and applications (like intelligence puzzles or art applications) are progressively taking their share in the limitless panorama of m-applications. Besides, as the use of these mobile learning educational applications has become more common anywhere in the world, new concerns have appeared in the classroom, human interaction in software engineering and ergonomics. New tests of foreign languages for a number of purposes have become more and more

common recently. However, studies interrelating language tests and ergonomics have lagged behind to the point that there is a clear lack of balance between software for m-learning and the capacities of modern equipment.

According to these indicators, the researchers worked and experimented with 26 high school students in Valencia (Spain) towards the implementation of a mobile phone language testing system for the Spanish University Entrance Examination. Since, extensive research is currently being carried with the results of such experimentation, this paper is just a tentative approach to the observation based results of the major problems found in such experimentation. Thus, it is not our intention to present conclusive results but to draw a general picture of what aspects may conflict in designing this type of application. Our conclusions present suggestions for further use and research.

II. CONSTRAINTS IN TESTING THROUGH MOBILES

In order to accomplish our research, the project engineers designed a mobile phone emulator that students could recognize and analyze. Our heuristic did not require an important allowance to run correctly. Furthermore, HIE did not require an unproven refinement to run correctly. Figure 1 presents the diagram used by the research team.



Figure1. Interface emulator

Our evaluation represents a valuable research contribution in itself. Our overall performance analysis obtained three main results:

- Visual ergonomics* present significant differences in task elaboration. For instance, students evidenced more favorable responses to using mobiles for speaking and listening activities. Thus, writing and reading were considered extremely uncomfortable.
- Creation and deployment of questions* through visual iconicity was considered acceptable but the time required by the device to operate was considered extremely low especially if tasks had been created through random generation or the test had been based on adaptive algorithms. In a practical mode, this meant that the transition between one task and the following was considered slow.
- Keyboard use* could be considered null for longer questions (like summaries or writing compositions) but the system was considered optimal for multiple choice tasks and speaking activities.

A. Visual ergonomics

Visual ergonomics has been considered a long standing issue in computer assisted language learning. However, if this is so in general learning, in testing becomes a crucial matter since an inadequate presentation of contents can challenge the validity of the test [1]. García Laborda [2] and García Laborda et al. [3] suggest that visual ergonomics play a significant role in the final result of the test. Thus interfaces in mobile learning should be simplified in the following aspects: colors, font sizes, lining, alignment and interline indenting [4]. Besides, buttons may be used but keyboard is preferable. Most of these guidelines which are often used for computer interfaces are especially crucial in mobile learning. Since the student will not have much visibility of written or even question texts, it could be advisable to use just one question at any one time and, whenever possible, limit the space of response to what can be seen at a glance.

B. Creation and deployment of questions

It is not only the setup what needs to be transformed. Questions and the delivery processes also need to change in order to make language tests accessible. Since we mentioned that questions need to be fully seen at first sight, testing items will necessarily be of limited space and, whenever possible, reduce the alternative points, bullets or responses to one line. This is however, a point that has not been adequately emphasized. Just by looking at figure 1, we can assume that it will almost always be like that but in reality there is still an open debate on how the construct and rubrics of the questions really affect the test performance. As in regular computers the scrolling process has been argued for a long time as a discriminating factor in language testing (apart from a factor that should be avoided by principle) [6]. However, although reading may be problematic, mobile devices may be useful to train towards listening and speaking tests [7]. Not only so but, the fact that students are flexible in the use of mobile devices for test training provides opportunities for more rehearsing and facilitates the acquisition of correct pronunciation patterns.



Figure 2. Types of questions (<http://hello-hello.com/blog/top-selling-ipad-app-developers-launch-language-learning-apps-for-the-iphone/>)

C. Keyboard use

One of the significant issues for years has been whether the fact that students need to use different tools in a computer based test than from a pen-and-paper test may lead to biasing the results. Today, it seems hard to understand of any single student in many parts of the world who cannot use a computer keyboard. However, results may vary according to its use. But it seems contradictory to accept that students may challenge their grades in computer based tests (especially older students) due to a possible lack of skill. Besides, keyboards for mobiles are very unpractical for writing compositions. As a consequence, one of the most significant drawbacks of the potential of mobiles for language testing is that we may not be able to use them to assess all four traditional skills.

III. EDUCATION IMPLICATIONS

Mobiles may prove a valuable asset for test training. Unless we change the way we assess languages (and still, writing is probably the most regarded skill), mobiles may have a limited use in real testing. They may be well acceptable for speaking and listening tasks. They may be also valid in situations in which a large number of students need to be evaluated through multiple choice items. However, as today mobiles have potential benefits as training tools. Only recently, some textbooks especially in English for Specific Purposes have begun to include some applications (apps) to support the students' self-manageable learning.



Figure 3. Interface of an app for learning Spanish (<http://hello-hello.com/blog/top-selling-ipad-app-developers-launch-language-learning-apps-for-the-iphone/>)

IV. CONCLUSIONS

In conclusion, we proved that despite the inherent complexity in designing a mobile learning device for language testing, our system could be used for parts in language testing that require extensive human use and thus reduce or simplify the testing process with an interesting reduction of economic cost by the test administrators and, what is even better, a reduced payment for the test candidate. Besides, since the ludic aspect of using mobiles makes them attractive for potential users, it could also be used to create training applications for prospective candidates who, in turn, would be able to use their testing skills in a comfortable and ever accessible context (in

relation to time and space). In conclusion, our model for language testing could be considered for its profitability in the short run. In fact, the main contribution of our work is that we confirmed that m-testing can reduce costs and be highly motivating for candidates. Very good examples in this sense can be found at <http://hello-hello.com/blog/>.

ACKNOWLEDGMENT

The authors would like to express their gratitude to the Ministry of Science and Innovation for providing the funds for this research through the OPENPAU project (ORIENTACION, PROPUESTAS Y ENSEÑANZA PARA LA SECCION DE INGLES EN LA PRUEBA DE ACCESO A LA UNIVERSIDAD; FFI2011-22442) thanks to the co-financing between the MICINN and FEDER funds in fulfillment of article 46 of the regulations (CE) n.º 1260/1999, of 21 June, of Consejo, y al Reglamento (CE) 1159/2000, of 30 May of the commission about information and publicity activities.

- [1] Y. Sawaki, "Comparability of conventional and computerized tests of reading in a second language", *Language Learning & Technology*, vol. 5, no. 2, 2001, pp. 38-59.
- [2] J. Garcia Laborda, "From fulcher to PLEVALEX: Issues in interface design, validity and reliability in internet based language testing", *CALL EJ*, vol. 9(1), 2007, no pagination, Retrieved July 1, 2011 at <http://callegj.org/journal/9-1/laborda.html>
- [3] J. Garcia Laborda, T. Magal-Royo, J. M. de Siqueira Rocha, and M. Fernández Alvarez "Ergonomics factors in english as a foreign language testing: The case of PLEVALEX" *Comput & Educ*, vol. 54, no.2, 2010, pp.384-391.
- [4] Hamel, Marie-Josée and Catherine Caws, "Usability tests in CALL development: Pilot studies in the context of the dire autrement and francotoile projects", *CALICO Journal*, 27, no. 3, 2010, Retrieved July 1, 2011 at <https://calico.org/a-814-Usability%20Tests%20in%20CALL%20Development%20Pilot%20Studies%20in%20the%20Context%20of%20the%20Dire%20autrement%20and%20Francotoile%20Projects.html>
- [5] T. Heift, "Drag or type, but don't click: A study on the effectiveness of different CALL exercise types", *Canadian Journal of Applied Linguistics/Revue Canadienne De Linguistique Appliquee*, vol. 6, no. 1, 2003, p. 69-85.
- [6] C. Parod-Ballester and Rodríguez, J. C., "Developing Spanish online readings using design-based research", *CALICO Journal*, 27, no. 3, 2010, 540-553. Retrieved July 1, 2011 at <https://calico.org/a-818-Developing%20Spanish%20Online%20Readings%20Using%20Design%20Based%20Research.html>
- [7] C. McBride, C. "Adaptive and maladaptive strategy use in computer-assisted language learning activities for listening comprehension", *Indian Journal of Applied Linguistics*, vol. 34, no.(1-2), p. 65-86. Retrieved July 1, 2011 at http://bahripublications.in/appliedlinguistic/back_volume.htm